Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 4 / ISSUE 3 / DECEMBER 2009 / 43-57
© 2009 Universidad de Talca - Chile

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762009000300005

Preference Analysis and Default Optimization in Web-based Product Configuration Systems

Reinhold Decker[1] and Sören W. Scholz[2]

Bielefeld University, Business Administration and Economics
[1] rdecker@wiwi.uni-bielefeld.de,[2] sscholz@wiwi.uni-bielefeld.de

Received 16 October 2008; received in revised form 18 May 2009; accepted 30 June 2009

Abstract

Today, more and more companies are providing web-based product configuration systems in order to better meet individual customer preferences. In many cases, pre-defined product specifications are additionally offered to facilitate the corresponding choice decisions. Against this background, we present a Poisson regression approach for analyzing customer preferences and a genetic algorithm for determining preference-based default products. The basis for this is transaction data, as they are automatically generated when configuring a product online. The potentials of the suggested methodology are demonstrated by means of two case studies referring to different product categories.

Key words: Customization, Genetic Algorithm, Poisson Regression, Preference Analysis, Product Configuration, Web Usage Data

1 Introduction

The diversity of products offered to meet customer needs has rapidly increased in recent years. Even simple devices with few basic functions, such as mobile music players or phones, nowadays provide a multitude of additional functions like USB interface, games, Bluetooth, etc. This development is accelerated by several factors, particularly the increasing globalization of competition, the availability of new technologies, and more demanding customers [28]. In addition, the Internet allows consumers an easy access to a wide range of information on product varieties, prices, conditions of supply, etc., which facilitates comprehensive overviews of alternative purchase options [37]. The Internet not only offers the opportunity to seek information but also provides product recommendations, which significantly improve the possibilities of identifying those products that adequately satisfy consumers' individual needs. At the same time, many industries are using one-to-one marketing techniques to better keep pace with the growing market dynamics and the increasing differentiation of individual preferences [40].

Nevertheless, the more products a manufacturer or retailer offers, the harder it is for the consumers to identify those which best meet their individual preferences. Studies have shown that consumers can even become overwhelmed if too many choices are offered and/or if those choices are poorly organized [17]. Particularly two options for supporting customers' purchase decisions are worth a closer examination in online business, namely the provision of passive search tools and that of active search tools [4]. Passive search tools, such as most recommender systems, use the consumers' revealed preferences from past purchases to suggest products that might be interesting in the current decision context. Active search tools, such as product configuration systems, allow customers to successively specify the products that presumptively best meet their individual preferences. The focus of this paper is on the latter.

According to [37], interactive environments like the World Wide Web are predestined for product and service customization. The net payoff of interactivity to consumers is usually assumed to be positive. Existing web-based configuration systems enable the individual customization of package tours (see, e.g., Site 1), laptops (see, e.g., Site 2), and even bags (see, e.g., Site 3) or sports shoes (see, e.g., Site 4). Car manufacturers, for instance, mostly offer large numbers of vehicle colors, equipment components, financing options, etc. However, retailers who are going to implement a high-variety strategy need to ensure that the customers are not confused by the complexity inherent in a wide range of options [16].

Recommendations provided by retailers or manufacturers are a confirmed means to concretize unclear customer preferences in the purchase decision process. Senecal and Nantel [35] p. 159, for example, showed by experimental research that consumers "who consulted product recommendations selected recommended products twice as often as subjects who did not consult recommendations". If the customer does not explicitly specify otherwise, he or she often automatically receives the "default option" [3]. According to [29], a "default" can be defined as an initial product design from which a customer can perform additional customization to account for his or her individual preferences. If defaults that are offered match existing consumer preferences, the above-mentioned complexity significantly decreases [7].

So, why do defaults matter in online product customization? First, they provide an implicit endorsement [22]. Second, they serve as an anchor point for choice decisions [20]. Finally, and importantly, a default may even enable an increase of turnover if its choice involves the implicit acceptance of a high-value product feature instead of its low-value counterpart (e.g., selection of a leather-trimmed steering wheel instead of a standard steering wheel in the case of the interior features of a car). In the latter case, the defaults should be defined in such a way that not only the consumer preferences are taken into account, but also the contribution margin of individual product features. Thus, optimal default products integrate consumer preferences and managerial objectives, which suggest a two-step procedure:

1. Measuring consumer preferences and definition of one or more initial defaults
2. Adapting preference-based defaults to given profit maximization goals

Both aspects will be addressed in the following. Therefore, the remainder of the paper is structured as follows: Section 2 briefly reviews alternative approaches for preference measurement using consumer surveys or web-based transaction data. In Section 3, a Poisson regression model is introduced which enables the measuring of consumer preferences in highly differentiated markets. Furthermore, a heuristic optimization approach based on a genetic algorithm is used to adapt preference-based initial default product candidates to external managerial and technical constraints. Section 4 presents two case studies focusing on areas where online configuration has gained special importance. By this means, we demonstrate the basic functionality of the suggested methodology. The paper concludes with a short discussion in Section 5 and a summary with an outlook on future research in Section 6.

2 Preference Measurement Using Purchase Data

Adequate specifications of customer-oriented defaults require reliable knowledge about the existing consumer preferences and behavior. This raises the question as to how this knowledge can be provided. A variety of preference measurement techniques have been developed in the last three decades. These methods can be subdivided with respect to the data used for measuring consumer preferences: Survey-based approaches and purchase-based approaches.

In general, survey-based approaches rely on interviewing respondents. Besides self-explicated approaches, which are based on the direct surveying of preferences for individual product characteristics, conjoint analysis has become the most widespread class of methods to measure consumer preferences in questionnaires [32]. Conjoint analysis decomposes holistic product evaluations or choices in metric part-worth utilities for the different attributes that make up the considered product [12]-[13]. In recent years, particularly choice-based conjoint analysis has established itself as an important preference measurement tool in marketing research and practice [33]. In addition, hybrid approaches, which combine self-explicated and conjoint approaches, are often used to measure preferences when a great number of product attributes have to be taken into account [23], [26]. Moreover, multi-criteria decision support methods have been adapted to measure preferences in consumer surveys. Particularly, the Analytic Hierarchy Process (AHP) has been successfully applied to measure consumer preferences for complex products [23], [34].

While survey-based preference measurement approaches have some undisputed advantages–especially when really new product ideas have to be assessed–there are some well-recognized drawbacks: Surveys usually require the cooperative participation of the respondents/consumers to identify relevant information about their predispositions or consumer behavior. However, return rates of interviews have been declining in the last three decades [6]. This effect can lead to devastating consequences for the quality of the obtained preference estimates. For example, Simonson [36], p. 36 states: "Specifically, customers who believe that they have strong, well-defined, and informed preferences are likely to place greater value on their participation in the process than are customers who are less sure that they know what they want." Accordingly, in recent decades, the budget for empirical surveys has further increased to compensate for this effect [5]. Currently, best practices demand that researchers should try to maximize response rates to decrease the risk of nonresponse errors to a minimum. Unfortunately, recent literature on this topic shows that the elimination of nonresponse errors, e.g., by means of repeatedly inviting the respondents, often results in an increase of measurement errors [25]. One possible reason for this may lie in the lower interest of these laggard subjects to participate, which may affect the quality of their answers.

Using real purchases to measure consumer preferences provides a valuable alternative to using hypothetical questioning in consumer surveys. Today, the analysis of consumer behavior is facilitated by the fact that more and more purchases are made on the Internet. The respective transaction data can be used to analyze preferences almost cost-free. An early approach for mining customer preferences using web log data was suggested by [15]. However, this type of approach basically differs from the following one in that it does not explicitly measure preferences but primarily anticipates individual preferences by considering similarities of purchasing patterns.

As already indicated in Section 1, online configuration systems are promising sources for providing such data. Rusmevichientong et al. [31], p. 45, for instance, note that the web-based GM Auto Choice Advisor provides “access to large quantities of data that reflect consumer preferences”, whereas [9] point out that to identify and extract actionable information is still a bottleneck in the analysis of these data.

In their renowned paper, [11] applied a discrete choice model to POS scanner data in a conjoint-like manner in order to estimate consumer preferences for distinctive product features. Their approach provides parameter estimates that can be interpreted similarly to the part-worth utilities known from survey-based conjoint analysis [12], [33].

Therefore, at a first glance, discrete choice models seem to be promising vehicles to extract consumer preferences from online purchase decisions. However, these models assume that consumers choose between a predefined set of different products, which implies that the individual consideration sets are known and sufficiently small. Swait and Ben-Akiva were probably the first to examine the impacts of a misspecification of the consideration set in discrete choice analysis [39]. Therefore, various two-stage approaches have been proposed that aim to identify the consumer’s consideration set before applying a discrete choice model (see, e.g., [2], [38]). These two-stage approaches have been proven successful in capturing non-compensatory behavior in decision problems including small numbers of alternatives, but they suffer from the combinatorial complexity arising from the latent nature of the consideration set formation process [2]. This hampers the application of these models to purchase situations where hundreds, or even thousands, of different product alternatives are available.

In the case of traditional conjoint analysis, [18] addressed the substantial error that emerges when the consideration set of a consumer is not identical with the set of products available. Since many product configuration systems explicitly enable the definition and purchase of a large number of different alternatives within one product category, the validity of conjoint analysis and discrete choice models can be assumed to be rather poor when analyzing such online markets.

The method outlined in Section 3 explicitly tries to account for the existence of large numbers of different product alternatives by using a Poisson regression approach. The Poisson distribution is typically used to model the occurrence of rare events within a given time period. In the present context, the “rareness” of a particular alternative is a direct (and in principle even intended) effect of applying an online product configurator. Using such a system for determining the optimal, or at least a satisfactory product, implicates a different decision logic as it underlies the choice between a given set of alternatives [19], [30]. Individually specifying a product by means of a configuration system structurally equals a sequential decision process where one attribute is updated after the other.

3 Methodology

The method presented in this paper consists of two parts: First, a Poisson regression model is applied to measure consumer preferences based on purchase data collected in online log-files. Based on the measured preferences, a genetic algorithm is used in the second step to identify an optimal assortment of default products with respect to a given criterion. Here, without loss of generality, the (overall) profit and the dissimilarity of a default product assortment are selected as target variables, because these criteria are typically considered as highly important performance figures by most vendors. The two parts of our approach are presented in detail in the following.

3.1 Preference Measurement by Means of a Poisson Regression Model

The following model is based on three main assumptions:

(1) A product can be described by a fixed number of attributes and is therefore represented by a specific combination of attribute levels (“features”).

(2) The purchase of a certain product substantiates the basic acceptance of the respective combination of attribute levels.

(3) The considered product category features a great variety of individual products.

Assumption (1) directly refers to the generally accepted understanding of how a product can be represented in conjoint analysis. Due to the fact that consumers normally have multiple options for satisfying their individual needs, it seems justified to assume the relation given by (2). Assumption (3) refers to the motivations given in connection with the increasing popularity of web-based product configuration systems in Section 1.

Let $k=1,\dots,K$ be the subscript used to identify different products, i.e., combinations of attribute levels, and $l=1,\dots,L$ the subscript for product attributes. Accordingly, K equals the number of different product profiles considered in a data set, each one chosen at least once. Then, with $m_l=1,\dots,M_l$ denoting the respective levels of attribute l, the consumers’ online purchases can be coded as follows:

$$(1) \qquad x_{klm_l}=\begin{cases}1, & \text{if attribute level } m_l \text{ of attribute } l \text{ is chosen with product } k\\ 0, & \text{otherwise}\end{cases}$$

If the frequency of purchasing a particular product is interpreted as the result of a counting process, with Y being the corresponding random variable, the Poisson probability of observing $y\in\{0,1,2,\dots\}$ purchases within a given period is

$$(2) \qquad \text{Prob}(Y=y)=\frac{\lambda^y}{y!}\exp(-\lambda)\,,$$

where λ equals the mean purchase rate in the product category of interest. $\lambda=2.5$, e.g., indicates that the product is purchased 2.5 times on average.

Assuming that purchase rate λ directly corresponds to the product specifications, and by referring to a similar approach by [10] for designing commercial websites, we suggest the following reparameterization:

$$(3) \qquad \lambda=\exp\left(\beta_0+\sum_{l=1}^{L}\sum_{m_l=1}^{M_l}\beta_{lm_l}x_{lm_l}\right),$$

where parameter $\beta_{lm_l}\in\mathbb{R}$ (with $l=1,\dots,L$ and $m_l=1,\dots,M_l$) represents the strength and direction of the relationship between attribute level m_l and purchase rate λ. The intercept parameter β_0 marks the corresponding baseline for the category as such. By using an exponential function, the non-negativity of λ is guaranteed. To simplify the notation, let $\boldsymbol{\beta}=(\beta_0,\beta_{11},\dots,\beta_{lm_l},\dots,\beta_{LM_L})\in\mathbb{R}^{\Sigma M_l+1}$.

In order to estimate the unknown parameters of the above Poisson regression model, the Maximum-Likelihood method can be used. The respective log-likelihood function reads:

$$(4) \qquad LL(\boldsymbol{\beta}) = \ln\left(\prod_{k=1}^{K} \frac{\lambda^{y_k}}{y_k!}\exp(-\lambda)\right) = \sum_{k=1}^{K} y_k \ln(\lambda) - \lambda - \ln(y_k!)$$

with $\lambda = \exp(\beta_0 + \sum_l \sum_{m_l} \beta_{lm_l} x_{klm_l})$. Because of the dichotomous definition of the product attributes, the parameter estimates $\hat{\beta}_{11},\ldots,\hat{\beta}_{LM_L}$ can be interpreted in a similar way to those resulting from conjoint analysis. This, in a broader sense, allows us to call them part-worth utility-like coefficients. Accordingly, $\hat{\beta}_{21}$, for example, corresponds to the estimated relation between the first level of attribute 2 and the purchase probability of a product. In a real setting, where attribute 2, for example, is the surface finish of a laptop with level 1 equaling “piano black”, $\hat{\beta}_{21} = 0.15$ would indicate the positive influence of a piano black finish on the purchase probability of the respective laptop, whereas $\hat{\beta}_{24} = -0.30$ for a green finish would point in the opposite direction with greater strength.

3.2 Constrained Optimization of Default Product Assortments

Despite the increasing popularity of product configuration systems, it cannot be taken for granted that all consumers actually prefer to customize their individual products in a “have it your way” approach [21]. Depending on the considered product category, the configuration task can become quite tedious and time-consuming. Large diversities of options (i.e., attributes and attribute levels in the present context) can even discourage the customer in the case of limited experience in customizing products or may end up with an unfavorable confusion of the customer [14], [16].

In order to reduce the complexity of the customization process, retailers or manufacturers may additionally offer small numbers of predefined products (defaults) which presumptively satisfy different customer preferences. In doing so, they can prevent possible aversion to the use of configuration systems. Default products may also provide a substantial surplus for those customers who are willing to use a configuration system if they are provided as a starting point for the subsequent customization process.

Before discussing details of the optimization process, we will introduce some notation: p, with $p = 1,\ldots,P << K$, is the subscript for identifying different default products, $\mathbf{x}_k = (x_{k11},\ldots,x_{klm_l},\ldots,x_{kLM_L}) \in \{0,1\}^{\Sigma M_l}$ is a binary vector which represents an observed product profile k, and $\mathbf{X} = \{\mathbf{x}_1,\ldots,\mathbf{x}_k,\ldots,\mathbf{x}_K\}$ is the set of product profiles being considered in empirical response analysis. The number of default products is assumed to be specified a priori by the marketing management. Furthermore, let $\hat{\lambda}_p = \exp(\hat{\beta}_0 + \sum_{l=1}^{L} \sum_{m_l=1}^{M_l} \hat{\beta}_{lm_l} x_{plm_l})$ be the estimated response of the market to default product p. With this in mind, the determination of an “optimal” set of default products or profiles $\mathbf{X}^{(P)} = \{\mathbf{x}_1,\ldots,\mathbf{x}_p,\ldots,\mathbf{x}_P\}$ can be realized by referring to the following aspects:

Profit maximization: The set of default products $\mathbf{X}^{(P)}$ should positively contribute to the overall profit of the retailer/manufacturer. Therefore, we maximize the following profit function for the set of defaults by systematically varying the attribute levels, i.e., the values of $\mathbf{x}_p$:

$$(5) \qquad \text{Profit}(\mathbf{X}^{(P)}) = \sum_{p=1}^{P} \text{Profit}(\mathbf{x}_p) = \sum_{p=1}^{P} \hat{\lambda}_p \left(\text{Price}(\mathbf{x}_p) - \text{Cost}(\mathbf{x}_p)\right),$$

where $\text{Price}(\mathbf{x}_p)$ and $\text{Cost}(\mathbf{x}_p)$ denote the unit price and the unit cost of default product p. The latter is defined as:

$$(6) \qquad \text{Cost}(\mathbf{x}_p) = \sum_{l=1}^{L} \sum_{m_l}^{M_l} x_{plm_l} c_{lm_l} \quad \forall p,$$

with c_{lm_l} indicating the costs that arise when level m_l of attribute l is considered. Fixed costs are ignored since they are not relevant for the sales-oriented optimization considered here.

Diversity: Ideally, the default products should represent the whole range of possible product alternatives. This would increase the probability of the consumer finding an adequate default product and reduce the complexity of the choice process [9]. In the long run, the provision of an adequately diversified set of default products may prevent the retailer/manufacturer from being too susceptible to unexpected changes in consumer purchase behavior.

The diversity requirement can be included in the optimization process by means of a similarity concept. According to the popular Tanimoto coefficient, the similarity between two default products $\mathbf{x}_{p_i}$ and $\mathbf{x}_{p_j}$ is defined by the following ratio:

(7) $$\mathrm{Sim}(\mathbf{x}_{p_i},\mathbf{x}_{p_j}) = a/b$$

$$\text{with } a = \sum_{l=1}^{L}(\beta_l^{\max}-\beta_l^{\min})\cdot\sum_{m_l=1}^{M_l} x_{p_i l m_l}\cdot x_{p_j l m_l} \text{ and } b = \left(2\cdot\sum_{l}^{L}(\beta_l^{\max}-\beta_l^{\min})\right) - a$$

Here, $\beta_l^{\max}$ equals the highest and $\beta_l^{\min}$ equals the lowest coefficient of all levels of attribute l. The weighting factor $(\beta_l^{\max}-\beta_l^{\min})$ is used to take into account the varying importance of different product attributes (the greater the aforesaid difference, the higher the importance). The new similarity measure $\mathrm{Sim}(\mathbf{x}_{p_i},\mathbf{x}_{p_j})$ rates two default products p_i and p_j to be similar if they differ, at best, with respect to less important attributes. In turn, both are rated to be dissimilar if they show differing levels on important attributes. The average pairwise dissimilarity between all default products included in $\mathbf{X}^{(P)}$ determines the diversity of the default product assortment, i.e.:

(8) $$\mathrm{Dissim}(\mathbf{X}^{(P)}) = \frac{2}{P(P-1)}\sum_{i=1}^{P-1}\sum_{j=i+1}^{P}\left(1-\mathrm{Sim}(\mathbf{x}_i,\mathbf{x}_j)\right)$$

In order to come up with an adequate relation between profit maximization and the degree of diversity of the default products, the following convex combination of both objectives is maximized:

(9) $$f(\mathbf{X}^{(P)}) = \alpha\cdot(1-1/\mathrm{Profit}(\mathbf{X}^{(P)})) + (1-\alpha)\cdot\mathrm{Dissim}(\mathbf{X}^{(P)}) ,$$

with $\alpha\in[0,1]$ to be specified exogenously by, for example, the marketing manager. Setting $\alpha=1$ leads to pure profit maximization while $\alpha=0$ results in the most diversified ("dissimilar") range of default products.

Constraints: In practice, certain combinations of attribute levels might be prohibited or unwanted for technical, economic or legal reasons. In this case, profit maximization typically involves a more or less large number of constraints. Basically, at least three main types of constraints can be distinguished in the present context, namely technological restrictions, managerial requirements, and methodological conditions. Technological restrictions can result from the non-compatibility of certain attribute levels or features. For example, it is hardly advisable to combine a high-end video card with a low-end processor when configuring a gaming laptop. A related type of technological restriction concerns the non-admissibility of certain attribute levels. Furthermore, with regard to the completeness of the offered product range, the availability of a certain feature through at least one default might be required. In the laptop category, for example, it might be opportune to offer at least one default product that includes a digital TV tuner for those consumers who want to use their laptop as a television receiver as well. Finally, from a methodological point of view, one has to ensure that in each product profile $\mathbf{x}_p$ each attribute occurs with only one level. For instance, a laptop has either a 40 GB, an 80 GB, or a 120 GB hard drive. Accordingly, a typical set of constraints can look like this:

(10) $\mathrm{Sim}(\mathbf{x}_{p_i},\mathbf{x}_{p_j})<1 \quad \forall \mathbf{x}_{p_i},\mathbf{x}_{p_j}\in\mathbf{X}^{(P)} \quad \text{with } i\neq j$ (uniqueness of the elements of $\mathbf{X}^{(P)}$)

(11) $x_{pim_i}+x_{pjm_j}\leq 1 \quad \text{with } i\neq j, m_i, m_j \text{ fixed}$ (non-compatibility of attribute levels)

(12) $\sum_{p=1}^{P} x_{pim_i}\geq 1 \quad \text{with } i \text{ and } m_i \text{ fixed}$ (required availability of attribute levels)

(13) $\sum_{m_l}^{M_l} x_{plm_l}=1 \quad \forall p,l$ (uniqueness of attribute levels)

Depending on the considered product category and the market to be served, a great variety of constraints can be specified. Because of the conceptual character of this paper, we renounce a further discussion of this point.

After having specified both the objective function (Equation 9) and the set of constraints (according to the general examples given by Equations 10-13), the question arises how to solve the resulting constrained optimization problem. Due to the fact that the number of attributes and attribute levels, as well as the number of constraints, can easily become very large in real product configuration contexts, an efficient methodology is required.

3.3 Solving the Optimization Problem with a Genetic Algorithm

Genetic algorithms (GA) are powerful tools for solving non-linear optimization problems and are based on evolutionary concepts such as inheritance, mutation, and selection [24]. A common approach to deal with constraints within the GA framework is to use penalty functions that penalize infeasible solutions by reducing the value of the objective function (Equation 9) in proportion to the degree of constraint violation [42]. A simple and popular method is the so-called “death penalty” approach. Here, any infeasible solutions occurring during the optimization process are rejected without exception.

The basic idea of GA is to model the relevant search space by a set of individuals, each representing a possible solution to the given optimization problem. In the present context, an individual equals a set of default products $\{\mathbf{y}_1,\ldots,\mathbf{y}_P\}$, with $\mathbf{y}_p \in \mathbf{X}^{(P)}$. A set $\mathbf{Y}$ (with $\mathbf{Y} = \{\{\mathbf{y}_1^{(1)},\ldots,\mathbf{y}_P^{(1)}\},\ldots,\{\mathbf{y}_1^{(S)},\ldots,\mathbf{y}_P^{(S)}\}\}$) of S individuals is called a population. Subsequent populations are called generations. A simple way of generating S individuals that can be used as an initial population for the GA is to randomly select $S \cdot P$ existing product profiles from the data set underlying the whole optimization (i.e., from $\mathbf{X}$). In doing so, the basic admissibility of the starting solution is guaranteed.

In the present study, we applied an elitist GA, where the n (with $n \le S$) best individuals (the “elite”) of the current generation (the parents) are adopted for the next generation (the children) as they are. The corresponding reproduction probabilities are dynamically computed during the optimization process. Accordingly, the GA for the present optimization problem works as follows:

```
1. Randomly select $S$ individuals (sets of defaults) defining the starting population $\mathbf{Y}_0$ (i.e., evolution starts with $S$ parents) and initialize iteration counter $t = 1$.
2. Do until $t > T$
   2.1 Create $S$ children $(\tilde{\mathbf{Y}})$ by means of crossover and mutation.
   2.2 Create a new population $\mathbf{Y}_t$ by applying elite selection.
   2.3 Evaluate each individual $\{\mathbf{y}_1^{(s)},\ldots,\mathbf{y}_P^{(s)}\}$ of population $\mathbf{Y}_t$ according to (see Equation 9):
       $fit = \begin{cases} f(\{\mathbf{y}_1^{(s)},\ldots,\mathbf{y}_P^{(s)}\}) & \text{if } \{\mathbf{y}_1^{(s)},\ldots,\mathbf{y}_P^{(s)}\} \text{ is a valid option} \\ f(\{\mathbf{y}_1^{(s)},\ldots,\mathbf{y}_P^{(s)}\}) - \Delta & \text{otherwise} \end{cases}$
   2.4 Compute fitness proportional reproduction probabilities and increase $t$ by 1.
3. Choose that set of defaults $\{\mathbf{y}_1^{opt},\ldots,\mathbf{y}_P^{opt}\} \in \mathbf{Y}_T$ which yields the maximum fitness value $fit$.
```

Here, T is the maximum number of iterations and Δ denotes the penalty term. Both parameters have to be defined adequately (e.g., $T = 1{,}000$ and $\Delta = 10^{10}$).

4 Empirical Applications

In order to demonstrate the basic functionality of the suggested approach, two data sets are considered in the following. The first one was generated from log-file data provided by an online package tour operator, whereas the second data set refers to the offerings of an online shop for laptops. The first example was selected to illustrate the Poisson regression based modeling of the so-called Long Tail phenomenon: Nowadays, a substantial amount of the sales volume of many companies disperses over a wide range of products that partly only sell a few times each. Anderson called this phenomenon the “Long Tail” to refer to the shape of the respective sales distribution curve [1]. The phenomenon can be observed particularly in online markets. However, because of the individuality of holiday trips, optimal default packages interesting for a larger number of consumers are hard to determine. Therefore, additionally the laptop market is considered where vendors might tend to reduce the Long Tail effect in order to deal with the problem of high capital commitment due to the provision of large numbers of alternative product components or features.

4.1 Consumer Preferences for Online Package Tour Options

The basic purpose of the first study is to show how aggregate preferences can be estimated from product, or in this case, service configuration data and, as mentioned above, how a Poisson distribution can be used to represent the Long Tail phenomenon. The data set includes 3,561 tour bookings that were taken in 2004 and 2005. Each booking is represented by ten attributes featuring different numbers of levels (see Table 1 for details).

If the shares of products showing 1, 2, 3, and so on, online bookings are plotted, the typical Long Tail structure results. Figure 1 shows both the observed shares (indicated by squares) and the approximation (indicated by asterisks) by fitting a Poisson distribution to the data (curve interpolation was used for illustration purposes only).

Obviously, most product profiles have booking frequencies lower than 5, the typical consequence of individual product, or rather tour, customization.

Applying the Poisson regression model suggested in subsection 3.1 to the data at hand provides the parameters depicted in Figure 2. A similar representation of parameter estimates was already used by [11] for illustrating aggregate preferences patterns. The present bar chart results from the purchases of 430 different travel or product profiles. In order to facilitate interpretations, the parameters have been normalized such that the values belonging to one attribute add up to 0 in each case. Wald tests (with $p = 0.05$) indicate that for each attribute, at least two levels are significantly related to booking frequency. Thus, each of the ten attributes explicitly contributes to the explanation of the observed frequencies.

Table 1: Available tour characteristics

Attributes	Attribute levels
Flight distance	Long-haul, short-haul
Length of stay	1-4 days, 5-13 days, > 13 days
Price	Up to €55 per day, €55-80 per day, > €80 per day
Accommodation	Single room, double room, apartment
Catering	All-Inclusive, half board, breakfast only, self catering
Hotel category	Up to 2 stars, 3 stars, 4 stars, 5 stars
Distance to beach	Up to 100 m, 101-500 m, 501-1,000 m, > 1,000 m
Distance to downtown	Up to 1 km, 1-3 km, 3-6 km, > 6 km
Hotel size	Up to 50, 51-100, 101-250, 251-500, > 500 rooms
Climate	Tropical, subtropical

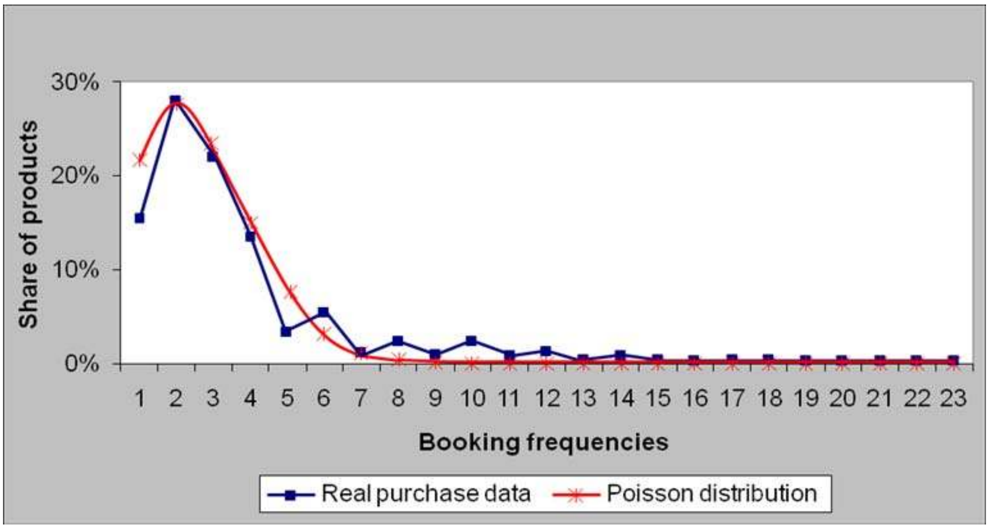

Figure 1: Long Tail pattern in the package tour market

As seen in Figure 2, the attribute “accommodation” is strongly related to the booking frequency. The fact that particularly double rooms and apartments obviously meet consumers’ preferences (indicated by the positive bars pointing to the right) is less astonishing when taking into account that the online travel agency considered here primarily offers holiday packages for families and couples who typically prefer these accommodation options. Single rooms, on the other hand, are rather an option for business travelers. But these travelers rarely book their business trips via package tour operators like the one considered in this study. So, the distinctive negative parameter is very plausible.

The fact that lower hotel categories (up to three stars) are seemingly favored over higher categories (four or five stars) may be due to the last-minute nature of parts of the package tour offerings. The parameters belonging to attribute “distance to downtown” tell us that the option of staying in an urban area (distance less than 1 km) has a positive influence on the booking frequency. The two middle options (1-3 km and 4-6 km) are less favored, while a positive effect also results from the attribute level “> 6 km”. This quite remarkable pattern can be explained to a certain degree by the fact that entertainment-oriented tourists opt for the closeness of a pulsating city and night life, whereas those tourists who are interested in tranquil recreation prefer the hinterland.

The plausibility of the regression results also finds its expression in the parameters that have been estimated for the attribute “length of stay”. Here, the clear preference for a medium length of the vacation trip is consistent with the fact that package tours are often considered for annual holidays. A coherent structure also exists for the attribute “climate”, where the moderate way of enjoying holidays in the warmth is preferred. Analogous interpretations are also possible for the remaining attributes, but have to be left to the reader for lack of space.

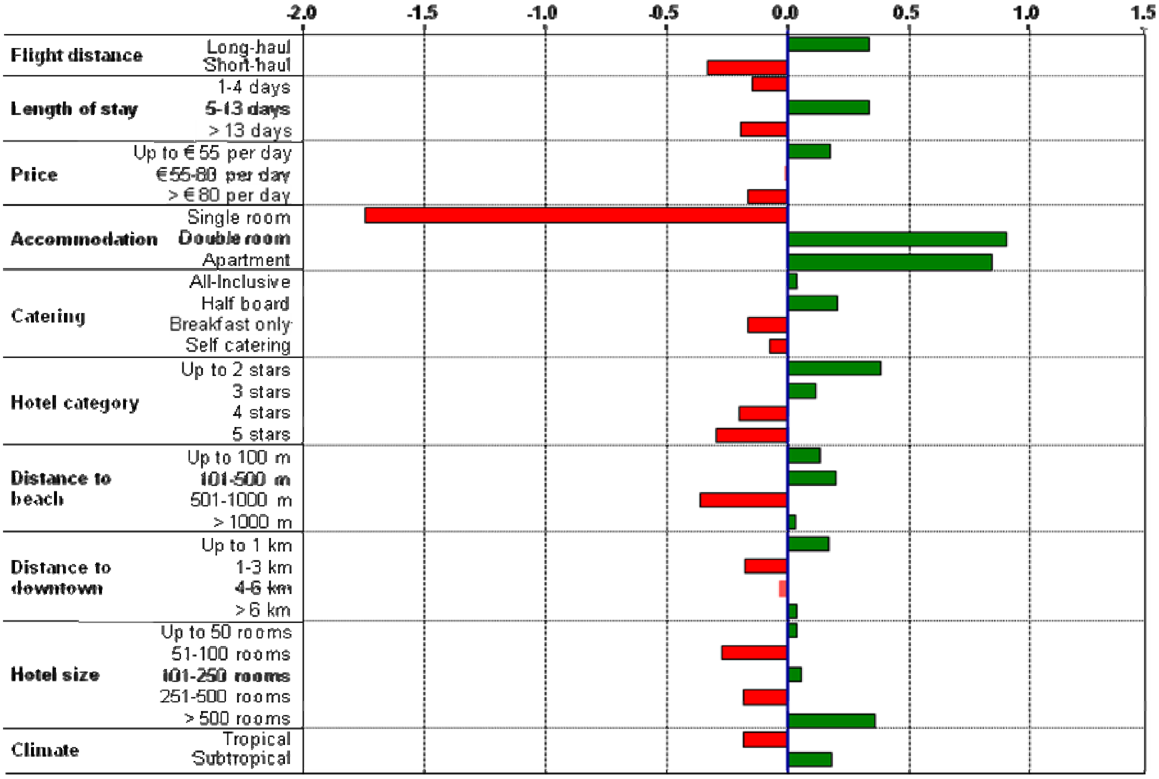

Figure 2: Parameter profile for package tour attributes

A Poisson regression not only allows us to measure aggregate preferences for relevant product or service features but also enables simulations that help to check the potential acceptance of alternative (possibly not yet available) offerings. A simple example is given in Table 2. The right-hand column shows a concrete package tour that would fit the profile given in column 2 (“level”). According to the model at hand, this option would imply 8.8 bookings in the sample considered, which significantly tops the observed average booking frequency of 4.05, and therefore constitutes an offer that could be worth considering as a default.

Table 2: Predicting the booking frequency of a new tour offer

Attribute	Level	Estimated regression parameter	Possible implementation
Flight distance	Short-haul	-0.33	Tuscany
Length of stay	1-4 days	-0.15	3 days
Price	> €80 per day	-0.17	€85 per day
Accommodation	Double room	0.84	Double room
Catering	Breakfast only	-0.17	Breakfast only
Hotel category	Up to 2 stars	0.38	2 stars
Distance to beach	501-1,000 m	-0.36	800 m
Distance to downtown	Up to 1 km	0.17	500 m
Hotel size	101-250 rooms	0.05	120 rooms
Climate	Subtropical	0.18	Subtropical
			Predicted booking frequency: 8.8

In order to avoid misinterpretation, it has to be mentioned that the selected example “Tuscany” only serves as an illustration of a short flight distance. The specific touristic features of Tuscany (or any other destination) are not included in the data used for model calibration and therefore have to be excluded from interpretation. Meaningful preference measurements on such a differentiated level are not possible when limiting the data generation process to attributes with a manageable number of levels.

4.2 Default Optimization in the Laptop Market

The second example refers to the laptop computer market, where product customization has gained increasing importance, as can be seen from the relevant activities of companies like Dell$^{\text{TM}}$ and Toshiba$^{\text{TM}}$. As already mentioned in Section 1, more and more vendors also provide decision support by offering pre-configured defaults. Against this background, the following example shows how optimal default laptops can be determined by considering consumer preferences as well as technological constraints and costs.

The data set used for illustration purposes includes $L = 15$ attributes with altogether 36 levels. Table 3 provides a structural description of the data. According to the configuration system that underlies our example, we excluded the screen size from consideration to ease the comparison of different configurations. Altogether, $K = 1{,}517$ different product profiles are considered in the following.

Table 3: Available laptop characteristics

Attributes [Abbreviation]	**Attribute levels**
Video card [VC]	NVidia 256 MB, NVidia 512 MB
Processor [CPU]	1.6GHz/1024KB, 1.8GHz/512KB, 1.8GHz/1024KB
Memory [RAM]	512MB-400MHz, 1024MB-400MHz
Hard drive [HD]	60 GB, 80 GB, 100 GB
Optical drive [OD]	DVD +/- RW Lightscribe, DVD +/- RW Standard
Integrated webcam [WC]	Not available, 1.3 MPixel webcam
Integrated wireless LAN [WL]	Not available, 802.11b/g (54MBit)
Bluetooth [B]	Not available, integrated Bluetooth adapter
TV tuner [TVT]	Not available, TV tuner (analog), TV tuner (digital/DVBT)
Operating system [OS]	No OS, Windows XP Home, Windows XP Professional
MS Office software [MSO]	Not available, Office 2003 Basic, Office 2003 Professional
Carry case [CC]	No carry case, standard carry case (19 inch)
Additional power pack [APP]	Not available, available
Car adapter [CA]	Not available, available
Mouse [M]	Pointing stick, touchpad, optical mini mouse

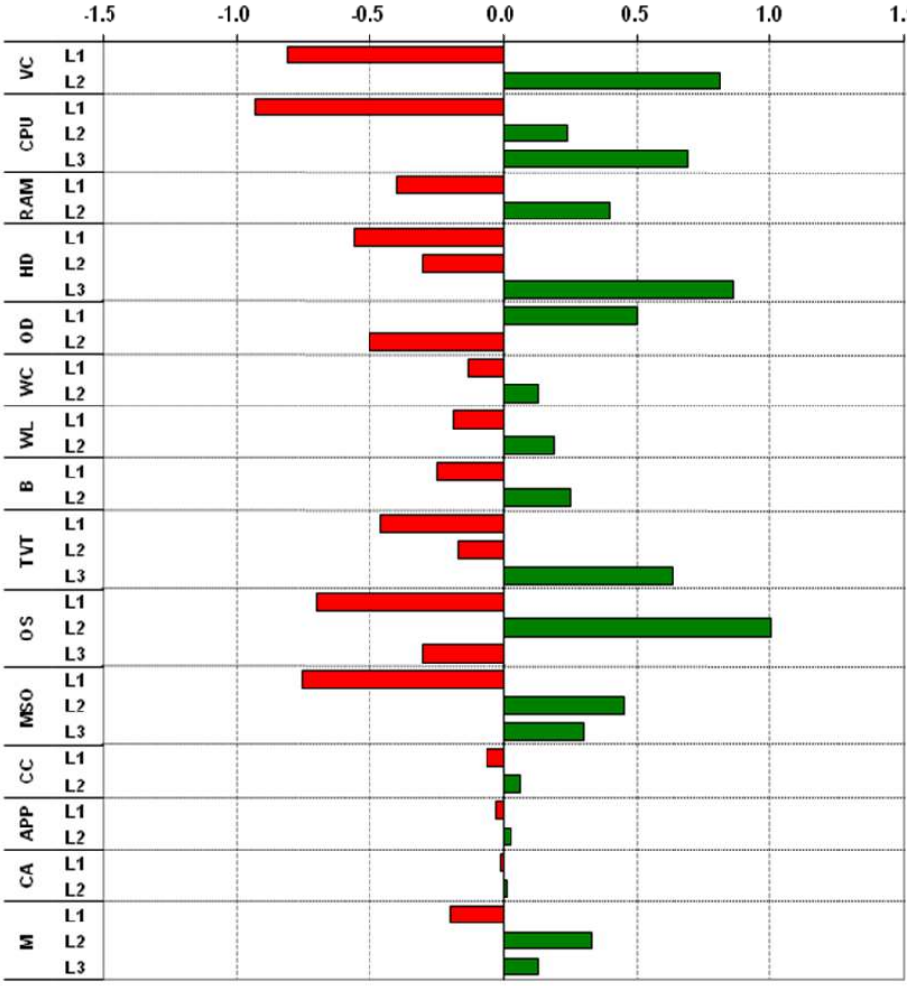

Figure 3: Parameter profile for laptop attributes

The parameters used in the optimization process are depicted in Figure 3. The abbreviations written in capital letters refer to the respective attribute (see Table 3), whereas L1, L2, and L3 refer to the related levels. The interpretation of the parameters is the same as with the package tour example. Once again, positive parameter values indicate positive preferences for the respective attribute levels and vice versa.

Furthermore, we assumed that the laptop vendor intends to offer $P = 5$ different default products satisfying consumer preferences measured by means of a Poisson regression. Because of the non-availability of the real costs of the relevant attribute levels (the crucial trade secret), we used plausible estimates based on the sale prices of the individual features taken from the product configurator.

Besides the methodological condition given by Equation 13, the following five constraints are exemplarily formulated for demonstration purposes and are taken into account by the suggested penalty function approach:

1. The 1.8GHz/1024KB processor (CPU) will not be offered with a car adapter (CA).
2. At most, one of the following three attribute levels may be installed in a default laptop: a small video card (VC), a small processor (CPU) or a small hard drive (HD).
3. Default laptops that are configured without an operating system (OS) are only offered without MS Office software (MSO).
4. The resulting set of defaults has to cover each level of the attribute “operating system” (OS).
5. All five product defaults have to be unique with respect to their individual combination of attribute levels.

The mathematical equivalents of these constraints read as follows:

(c1) $$x_{p,2,3} \cdot x_{p,14,2} = 0 \quad \forall \mathbf{x}_p \in \mathbf{X}^{(P)}$$

(c2) $x_{p,1,1} + x_{p,2,1} + x_{p,4,1} \leq 1 \quad \forall \mathbf{x}_p \in \mathbf{X}^{(P)}$

(c3) $x_{p,10,2} + x_{p,10,3} \geq x_{p,11,2} + x_{p,11,3} \quad \forall \mathbf{x}_p \in \mathbf{X}^{(P)}$

(c4) $\left(\sum_{p=1}^{P} x_{p,10,1}\right) \cdot \left(\sum_{p=1}^{P} x_{p,10,2}\right) \cdot \left(\sum_{p=1}^{P} x_{p,10,3}\right) \geq 1 \quad \forall \mathbf{x}_p \in \mathbf{X}^{(P)}$

(c5) See Equation 10

By using the objective function given with Equation 9, the GA determines optimal or near optimal solutions for varying α. Figure 4 shows the profit (represented by $\text{Profit}(\mathbf{X}^{(P)})$) and the diversity (represented by average pairwise dissimilarity $\text{Dissim}(\mathbf{X}^{(P)})$) of the resulting sets of default products when systematically decreasing α in steps of 0.1 from 1 to 0. For better readability, the profit has been normalized. Therefore, both the profit and the diversity only take values between 0 and 1.

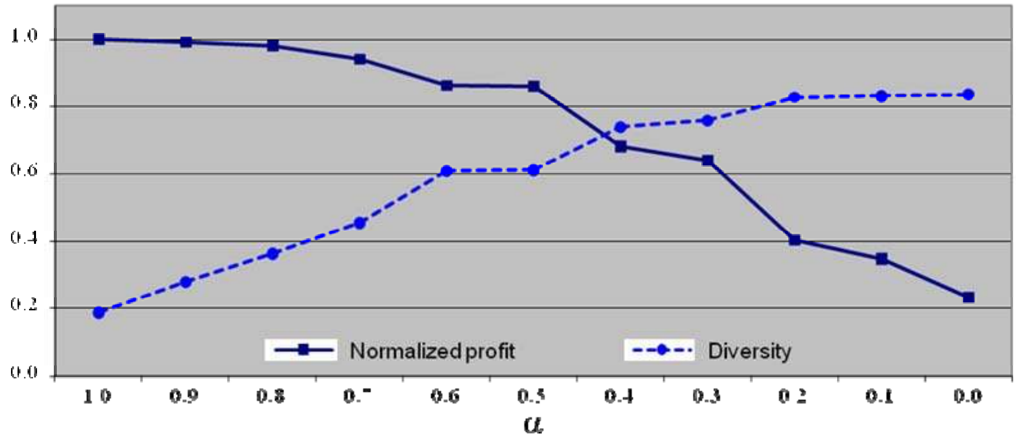

Figure 4: Normalized profit and average dissimilarity of optimal sets of default products

The final question as to which of these solutions is the "adequate" one is still up to the product manager. Here, of course, further information such as the default assortment of important competitors, the image positioning of the vendor, and the desired degree of diversity, also play a crucial role. However, the GA helps to limit the number of different solutions that have to be checked "manually" to a manageable size.

Table 4 exemplarily illustrates the optimal solution identified for $\alpha = 0.7$. The black bullets mark the respective attribute levels for the five default products. A closer consideration of the profiles shows that all constraints are taken into account. None of the profiles equals the "trivial" solution, which results when simply combining the attribute levels with highest parameter values. This, indeed, would maximize the user's benefit but may not necessarily fit the managerial intentions of the vendor.

Table 4: Optimal set of default products

p	VC		CPU			RAM		HD			OD		WC		WL		B		TVT			OS			MSO			CC		APP		CA		M		
1		●			●		●			●	●			●		●		●			●		●			●			●		●	●			●	
2		●			●		●			●	●			●		●		●			●			●			●		●		●	●			●	
3		●		●			●		●		●			●		●		●			●	●			●				●		●		●			●
4		●		●			●			●	●			●		●		●			●			●			●		●		●		●		●	
5		●			●		●		●		●			●		●		●			●		●			●			●		●	●			●	

The table also shows that the required dissimilarity between the individual default laptops is realized by particularly focusing on the important attributes, namely CPU, HD, OS, and MSO. The less important attributes are predominantly kept constant across the defaults. Product profile № 2 can be labeled the "high-end default" providing the top features on most attributes, whereas default № 3 is a laptop with rudimentary features (e.g., without an operating system and thus without any MS Office software). The present defaults would generate 2,289 sales in total and a total profit of $854,268, which corresponds to an average unit profit of $373.2 (see Table 5).

Already, this comparatively small optimization problem including "only" 15 attributes implies about $7.24 \cdot 10^{27}$ different possible solutions, or individuals in the GA terminology. This vividly shows the necessity of a powerful optimization technique like the one outlined in Section 3.3. Both a "manual" search of the optimal defaults and one by means of simple enumeration techniques would not be efficient, particularly when the number of constraints becomes large.

Table 5: Profits and sales of the default products

p	**Profit** [in $]	**Sales volume**	**Profit per unit** [in $]
1	209,449	564	371.36
2	210,391	553	380.45
3	79,871	232	344.27
4	161,611	414	390.36
5	192,946	526	366.82

5 Discussion

The empirical examples have shown that the new approach presented in this paper may yield useful insights in the preferences underlying consumer purchase decisions. As indicated by Figure 1, the Poisson distribution is well suited to fit online configuration data, which can be easily extracted from log-files. The Poisson regression, in turn, facilitates the decomposition of aggregate consumer preferences reflected in online purchase patterns. The resulting parameter profiles, if convincing and comprehensible, enable vendors to understand (or at least confirm) the latent preferences of their customers. This allows producers to improve their new product development processes and vendors to align their (default) product assortments with respect to actual preference structures.

The principle of default optimization has been successfully outlined in the laptop computer example. Here, the GA identified those attribute profiles that maximize preselected criteria, namely the overall profit and dissimilarity of the set of default products. Of course, the objective function can easily be adapted to other targets by exchanging these criteria. Moreover, the stated constraints just served as an example to illustrate the flexibility of the proposed algorithm. By redefining the constraints and/or adequately modifying the integrated penalty function, the GA can be adapted to other market or managerial situations. To that effect, our approach is applicable to various conditions occurring in modern online stores.

However, it is noteworthy to remind that the GA is a heuristic approach, which does not deterministically find the global optimum but tries to identify the best or near-best solution evolutionarily. Thus, it is advisable to run it several times with different starting populations in order to overcome the problem of running into local optima instead of obtaining the global one.

6 Summary and Outlook

The present paper has introduced a new approach to determining aggregate consumer preferences by analyzing transaction data provided by web-based product configuration systems. The suggested Poisson regression model enables the estimation of part-worth utility-like parameters that can be used to assess the relative importance of product attributes and their levels. However, in contrast to widespread conjoint analysis, costly consumer surveys are not required. The practical potential of the regression model in a Long Tail setting was demonstrated using a data example from tourism. By combining the modeling idea with an elitist genetic algorithm, we furthermore suggested an efficient approach for preference-based default optimization. By systematically optimizing the defaults to be presented in an online configuration system, for laptops or automobiles, for example, the retailer or manufacturer can facilitate the consumer's decision process and simultaneously ensure the implementation of own managerial objectives. Because of the parsimony of the response model with respect to parameterization and the widely approved efficiency of genetic algorithms, the suggested approach should be applicable to product categories of high diversity, too.

Of course, the new approach is not without its shortcomings. A crucial aspect is the data itself. Because of the basic functionality of commercial product configuration systems, the data collected in this way only describe the purchase decision as it becomes evident regarding the own range of products. The competitive environment remains unconsidered. Moreover, consumer preferences for really new product features that are not yet available on the market cannot be measured. For the default optimization task discussed in this paper, this limitation is justifiable, particularly when taking into account the fact that the required web usage data are available without additional survey costs. Thus, we rather consider the proposed approach as a supplement than a substitute to the survey-based preference measurement approaches briefly discussed in Section 2.

Another issue is preference heterogeneity. An obvious and simple way of dealing with this aspect is to use additional click-through data or to combine the web data with observed characteristics of existing customers (e.g., from the shipping department) to identify segments of homogeneous preferences. Future research in this context should also be concentrated on techniques that explicitly allow an integrated modeling of customer heterogeneity: for example, latent class Poisson regression [41] or Negative Binomial regression [27].

Websites List

Site 1: Expedia Travel: Airline Tickets, Hotels, Car Rental, Airfares, & Vacations
http://ww.expedia.com

Site 2: Dell Laptops, Desktop Computers, Monitors, Printers & PC Accessories
http://www.dell.com

Site 3: Timbuk2 Bags
http://www.timbuk2.com

Site 4: Adidas
http://www.adidas.com

References

[1] C. Anderson, The Long Tail: How Endless Choice is Creating Unlimited Demand, London: Random House, 2007.

[2] M. Ben-Akiva, and B. Boccara, Discrete choice models with latent choice sets, International Journal of Research in Marketing, vol. 12, no. 1, pp. 9-24, 1995.

[3] C. L. Brown, and A. Krischna, The skeptical shopper: A metacognitive account for the effects of defaults options on choice, Journal of Consumer Research, vol. 31, no. 3, pp. 529-539, 2004.

[4] E. Brynjolfsson, Y.J. Hu, and M.D. Smith, From niches to riches: Anatomy of the long tail, MIT Sloan Management Review, vol. 47, no. 4, pp. 67-71, 2006.

[5] R. Curtin, S. Presser, and E. Singer, Changes in telephone survey nonresponse over the past quarter century, Public Opinion Quarterly, vol. 69, no. 1, pp. 413-428, 2005.

[6] E. de Leeuw, and W. de Heer, Trends in household survey nonresponse: A longitudinal and international comparison, in Survey Nonresponse (D. A. Dillman, J. L. Eltinge, R. M. Groves, and R. J. A. Little, Eds.). New York: John Wiley & Sons, pp. 41-54, 2002.

[7] B. G. C. Dellaert, and S. Stremersch, Marketing mass-customized products: Striking a balance between utility and complexity, Journal of Marketing Research, vol. 42, no. 2, pp. 219-227, 2005.

[8] R. Dhar, Consumer preference for a no-choice option, Journal of Consumer Research, vol. 24, no. 2, pp. 215-231, 1997.

[9] V. Dhar, and A. Sundararajan, Customer interaction patterns in electronic commerce: Maximizing information liquidity for adaptive decision making, Stern School of Business, New York University, New York, IS-99-17, 1999.

[10] X. Drèze, and F. Zufryden, A web-based methodology for product design evaluation and optimisation, Journal of the Operational Research Society, vol. 49, pp. 1034-1043, 1998.

[11] P. S. Fader, and B. G. S. Hardie, Modeling consumer choice among SKUs, Journal of Marketing Research, vol. 33, no. 4, pp. 442-452, 1996.

[12] P. E. Green, A. M. Krieger, and Y. Wind, Thirty years of conjoint analysis: Reflections and prospects, Interfaces, vol. 31, pp. 56-73, 2001.

[13] J. R. Hauser, and V. Rao, Conjoint Analysis, Related Modeling, and Applications, in Marketing Research and Modeling: Progress and Prospects (Y. Wind and P. Green, Eds.). New York: Springer, 2003, pp. 141-168.

[14] M. G. Helander, and J. Jian, Electronic product development (ePD) for mass customization, in Proceedings of the Conference WWDU 2002, Berchtesgarden, 2002.

[15] S. Holland, M. Ester, and W. Kießling, Preference mining: A novel approach on mining user preferences for personalized applications, in Knowledge Discovery in Databases: PKDD 2003 (Lavrac, N., D. Gamberger, L. Todorovski, and H. Blockeel, Eds.). Berlin: Springer, 2003, pp. 204-216.

[16] C. Huffman, and B. E. Kahn, Variety for sale: Mass customization or mass confusion?, Journal of Retailing, vol. 74, no. 4, pp. 491-513, 1998.

[17] S. S. Iyengar, and M. R. Lepper, When choice is demotivating: Can one desire too much of a good thing?, Journal of Personality and Social Psychology, vol. 79, no. 6, pp. 995-1006, 2000.

[18] K. J. Jedidi, R. Kohli, and W. S. DeSarbo, Consideration sets in conjoint analysis, Journal of Marketing Research, vol. 33, no. 3, pp. 364-372, 1996.

[19] R. Johnson, B. Orme, and J. Pinnel, Simulating market preference with “build your own data”, in 2006 Sawtooth Software Conference Proceedings, Sequim, 2006, pp. 239-253.

[20] D. Kahneman, J. L. Knetsch, and R. Thaler, Experimental tests of the endowment effect and the coase theorem, Journal of Political Economy, vol. 98, no. 6, pp. 1325-1348, 1990.

[21] J. Liechty, V. Ramaswamy, and S. H. Cohen, Choice menus for mass customization: An experimental approach for analyzing customer demand with an application to a web-based information service, Journal of Marketing Research, vol. 38, no. 2, pp. 183-196, 2001.

[22] C. R. M. McKenzie, M. J. Liersch, and S.R. Finkelstein, Recommendations implicit in policy defaults, Psychological Science, vol. 17, no. 5, pp. 414-420, 2006.

[23] M. Meißner, S.W. Scholz, and R. Decker, AHP versus ACA – An empirical comparison, in Data Analysis, Machine Learning, and Applications (C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, Eds.). Berlin: Springer, 2008, pp. 447-454.

[24] M. Mitchell, An Introduction to Genetic Algorithms, Cambridge: The MIT Press, 1996.

[25] K. Olson, Survey participation, nonresponse bias, measurement error bias, and total bias, Public Opinion Quarterly, vol. 70, no. 5, pp. 737-758, 2006.

[26] M. E. Pullman, K. J. Dodson, and W. L. Moore, A comparison of conjoint methods when there are many attributes, Marketing Letters, vol. 10, no. 2, pp. 1-14, 1999.

[27] V. Ramaswamy, E. W. Anderson, and W.S. DeSarbo, A disaggregate negative binomial regression procedure for count data analysis, Management Science, vol. 40, no. 3, pp. 405-417, 1994.

[28] K. Ramdas, Managing product variety: An integrative review and research directions, Production and Operations Management, vol. 12, no. 1, pp. 79-101, 2003.

[29] T. Randall, C. Terwiesch, and K.T. Ulrich, User design of customized products, Marketing Science, vol. 26, no. 2, pp. 268-280, 2007.

[30] J. Rice, and D. G. Bakken, Estimating attribute level utilities from “design your own product” data-3, in 2006 Sawtooth Software Conference Proceedings, Sequim, 2006, pp. 229-238.

[31] P. Rusmevichientong, J. A. Salisbury, L. T. Truss, B. Van Roy, and P.W. Glynn, Opportunities and challenges in using online preference data for vehicle pricing: A case study at general motors, Journal of Revenue and Pricing Management, vol. 5, no. 1, pp. 45-61, 2006.

[32] H. Sattler, and S. Hensel-Börner, A comparison of conjoint measurement with self-explicated approaches, in Conjoint Measurement: Methods and Applications (A. Gustafsson, A. Herrmann and F. Huber, Eds.). Berlin: Springer, 2003, pp. 147-159.

[33] Sawtooth Software, Report on Conjoint Analysis Usage among Sawtooth Software Customers, 2005. [Online]. Available: http://www.sawtoothsoftware.com.

[34] S. W. Scholz, and R. Decker, Measuring the impact of wood species on consumer preferences for wooden furniture by means of the analytic hierarchy process, Forest Products Journal, vol. 57, no. 3, pp. 23-28, 2007.

[35] S. Senecal, and J. Nantel, The influence of online product recommendations on consumers’ online choices, Journal of Retailing, vol. 80, no. 2, pp. 159-169, 2004.

[36] I. Simonson, Determinants of customers’ responses to customized offers: Conceptual framework and presearch propositions, Journal of Marketing, vol. 69, no.1, pp. 32-45, 2005.

[37] J. H. Steckel, R. S. Winer, R. E. Bucklin, B. G. C. Dellaert, X. Drèze, G. Häubl, S. D. Jap, J. D. C. Little, T. Meyvis, A. L. Montgomery, and A. Rangaswamy, Choice in interactive environments, Marketing Letters, vol. 16, no. 3/4, pp. 309-320, 2005.

[38] J. Swait, Choice set generation within the generalized extreme value family of discrete choice models, Transportation Research B, vol. 35, no. 7, pp. 643-666, 2001.

[39] J. Swait and M. Ben-Akiva, An analysis of the effects of captivity on travel time and cost elasticities, in Annals of the 1985 International Conference on Travel Behavior, Noordwijk, Holland, 1986, pp. 113-128.

[40] E. A. von Hippel, and R. Katz, Shifting innovation to users via toolkits, Management Science, vol. 48, no. 7, pp. 821-833, 2002.

[41] M. Wedel, W. S. Desarbo, J. R. Bult, and V. Ramaswamy, A latent class poisson regression model for heterogeneous count data, Journal of Applied Econometrics, vol. 8, no. 4, pp. 397-411, 1993.

[42] Ö. Yeniay, Penalty function methods for constrained optimization with genetic algorithms, Mathematical and Computational Applications, vol. 10, no. 1, pp. 45-56, 2005.